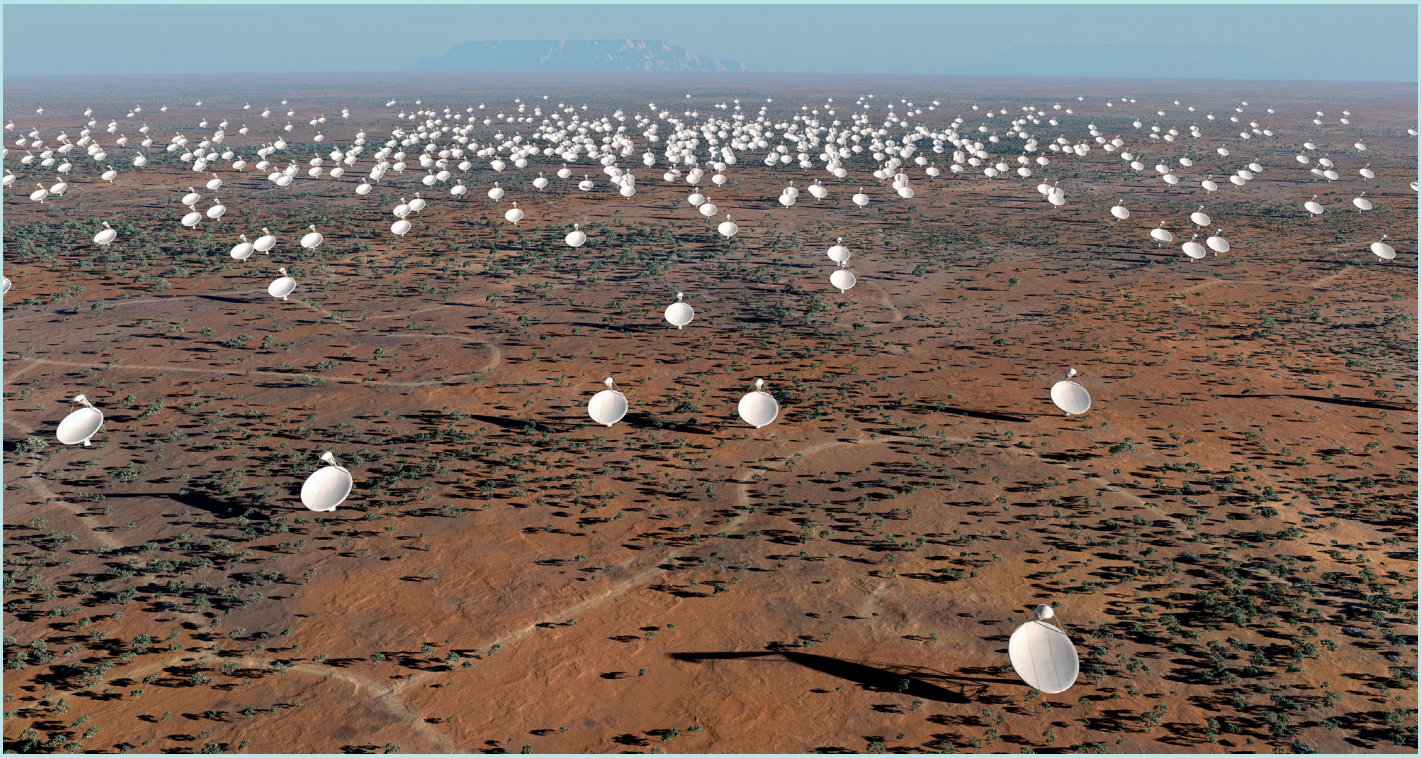
Questioni di *core*

Processori per i *big data*

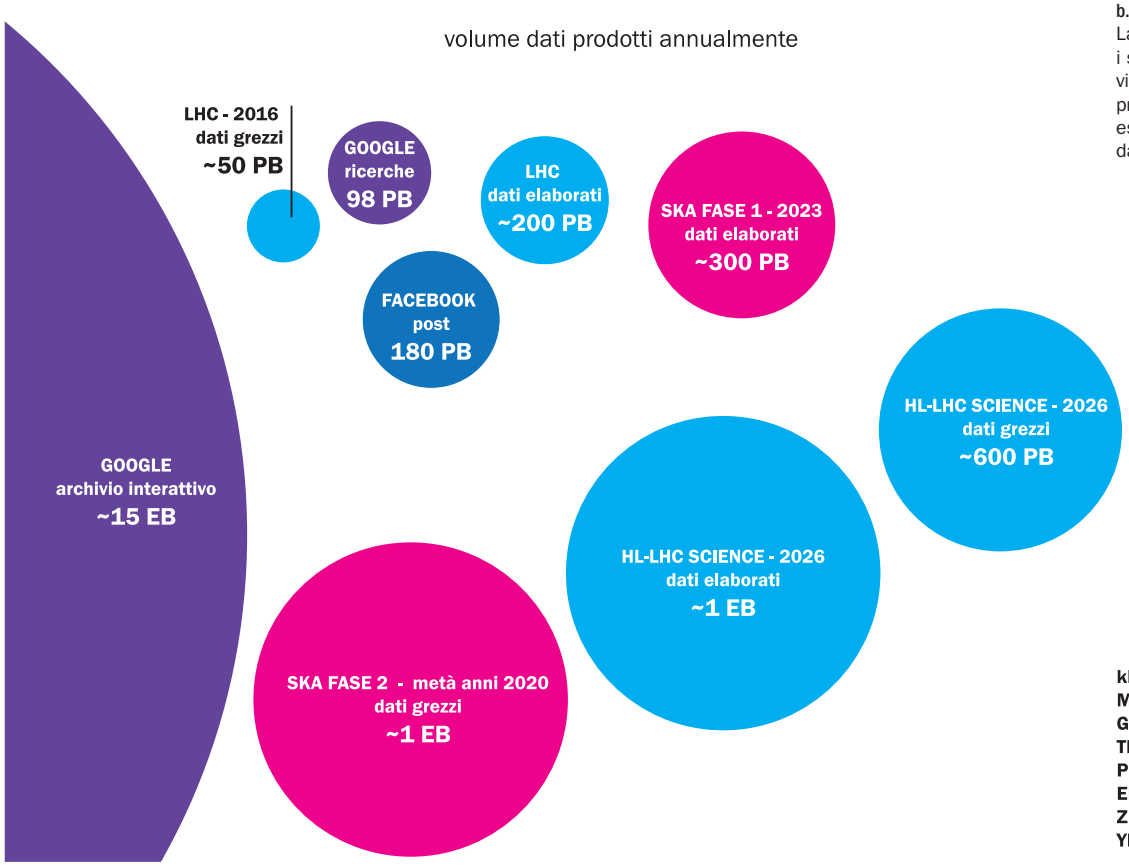
di Gaetano Maron



a.
Visualizzazione artistica del nucleo centrale, del diametro di circa 5 km, del radiotelescopio Square Kilometre Array (Ska). A regime il progetto prevede l'installazione di 130.000 antenne distribuite nei due siti in Sud Africa e Australia, che permetterà di avere una zona complessiva di raccolta delle onde radio provenienti dallo spazio di circa 1 km^2.

L'analisi dati nella fisica sperimentale, in particolare in quella delle alte energie, è oggi dominata, per dimensioni, potenza di calcolo necessaria e complessità degli algoritmi, dagli esperimenti del Large Hadron Collider (Lhc) del Cern. Anche la complessità e le dimensioni degli esperimenti di fisica astroparticellare e di astrofisica sta rapidamente evolvendo e la loro analisi dati presenta ormai caratteristiche simili a quelle di Lhc. Nel prossimo decennio, due giganti della *Big Science* inizieranno la presa dati: Hilumi-Lhc (Hl-Lhc), la fase 2 di Lhc, e Ska (Square Kilometre Array Telescope), il più grande radiotelescopio mai costruito al mondo, che sarà localizzato in Sud Africa e in Australia. Questo gruppo di esperimenti di nuova generazione produrrà dati al ritmo degli exabyte (vd. fig. b per il significato dei prefissi) all'anno, che equivale grosso modo alla capacità di *storage* attuale di tutti i centri di calcolo di Lhc distribuiti nel mondo. Anche l'analisi di quei dati richiederà potenze di calcolo ragguardevoli e, per il caso di Hilumi-Lhc, dell'ordine di 10 milioni di *core* di calcolo, circa 20 volte quelli disponibili attualmente per Lhc. Le Cpu (Central Processing Unit, v. anche p. 18, ndr) di un calcolatore, infatti, non sono più "monolitiche" (ossia costituite da un unico circuito integrato), ma contengono all'interno del proprio chip (o *"die"*) più unità di calcolo indipendenti tra loro che vengono chiamate *core*: una moderna Cpu può essere composta da qualche decina di *core*.
Queste risorse di calcolo pongono problemi sia tecnologici che di sostenibilità dei costi. La fig. b ci propone un interessante confronto, anche se solo indicativo, tra i dati prodotti da questi grandi esperimenti scientifici e quelli prodotti dai principali social network o dall'archivio Internet di Google. È alquanto impressionante notare come gli ordini di grandezza siano molto simili e che i casi scientifici abbiano a volte dimensioni anche maggiori. Nella figura si fa distinzione tra dati grezzi (in inglese *"raw data"*) e dati elaborati (in inglese

"*physics data*"). I primi sono i dati come vengono raccolti dai rivelatori dell'esperimento e contengono informazioni su quale elemento del rivelatore è stato colpito, sull'energia rilasciata, sul tempo, sulla posizione. Sono dati unici e non più riproducibili e per questa ragione ne vengono eseguite più copie preservate poi al Cern e negli altri centri di calcolo di Lhc. I dati elaborati derivano invece dal processamento dei dati grezzi, al fine di ricostruire l'impulso delle particelle, la loro origine, la loro energia, il tipo di particella rivelata, ecc., e contengono anche tutti i risultati finali prodotti dall'analisi dei canali di fisica studiati che possono portare a eventuali scoperte.
La *data science*, uno dei settori tecnologici a più elevato tasso di crescita, è nata sotto la spinta a creare sistemi efficienti di analisi dei *big data* dell'ordine di grandezza di quelli rappresentati in figura. Le tecnologie sviluppate in questo ambito utilizzano efficienti tecniche di memorizzazione e lettura dei dati, sofisticate tecniche statistiche per effettuare analisi predittive e, sempre più massicciamente, tecniche di intelligenza artificiale. Elemento comune a queste tecnologie è la possibilità di effettuare operazioni massicciamente parallele e avere architetture fortemente scalabili, anche a dimensioni elevate. L'analisi dati nella fisica delle alte energie utilizza già tecniche di *machine learning* (vd. p. 15, ndr) e sarà cruciale poter includere queste tecniche di "*big data analytics*" per far fronte alla complessità degli eventi di Hilumi-Lhc.
Uno degli aspetti che limitano la scalabilità dei software di analisi attuali, in particolare dei grandi esperimenti, è l'incapacità di sfruttare al meglio il parallelismo offerto dalle nuove generazioni di processori (*multi-core* Cpu, acceleratori massicciamente paralleli, *graphic processor*, ecc.). Per Hilumi-Lhc questo comporta una re-ingegnerizzazione del codice di analisi esistente, ottimizzandolo e integrandolo con le tecniche di analisi menzionate più sopra, spesso già predisposte al calcolo parallelo. Il software, inteso in senso lato, è quindi uno dei punti nevralgici dell'evoluzione dei sistemi di analisi verso l'era degli "exa-esperimenti", forse l'aspetto che più di altri avrà impatto sulle prestazioni e sulla scalabilità del sistema complessivo.
Evoluzione che nel settore dei processori spinge sempre più verso sistemi a elevato numero di *core* di calcolo per singolo Cpu, avendo ormai raggiunto limiti fisici alla potenza del singolo *core*: la famosa legge di Moore (secondo cui la potenza di calcolo del *core* raddoppia ogni 18/24 mesi), infatti, da qualche anno mostra evidenti segni di rallentamento. Questo approccio *multi-core* viene estremizzato nei *graphic processor*, sempre più utilizzati come acceleratori matematici in supporto alle Cpu convenzionali (vd. approfondimento). La tendenza sempre più marcata sarà quindi quella di sistemi ibridi formati da Cpu convenzionali, combinati con la potenza di calcolo straordinaria degli acceleratori.
Questo stesso approccio è seguito dai supercomputer di nuova generazione (detti anche "computer *exascale*", perché possono eseguire "exa", cioè miliardi di miliardi, di operazione al secondo). Questa comune linea di tendenza favorirà l'uso di queste gigantesche macchine di calcolo da parte della fisica sperimentale. È un approccio già in atto, che verrà consolidato in futuro e che potrà contribuire in modo significativo alle necessità



b.
La rivoluzione dei *big data*: i social e i motori di ricerca vi contribuiscono in maniera predominante, ma anche i grandi esperimenti scientifici agiscono da protagonisti.

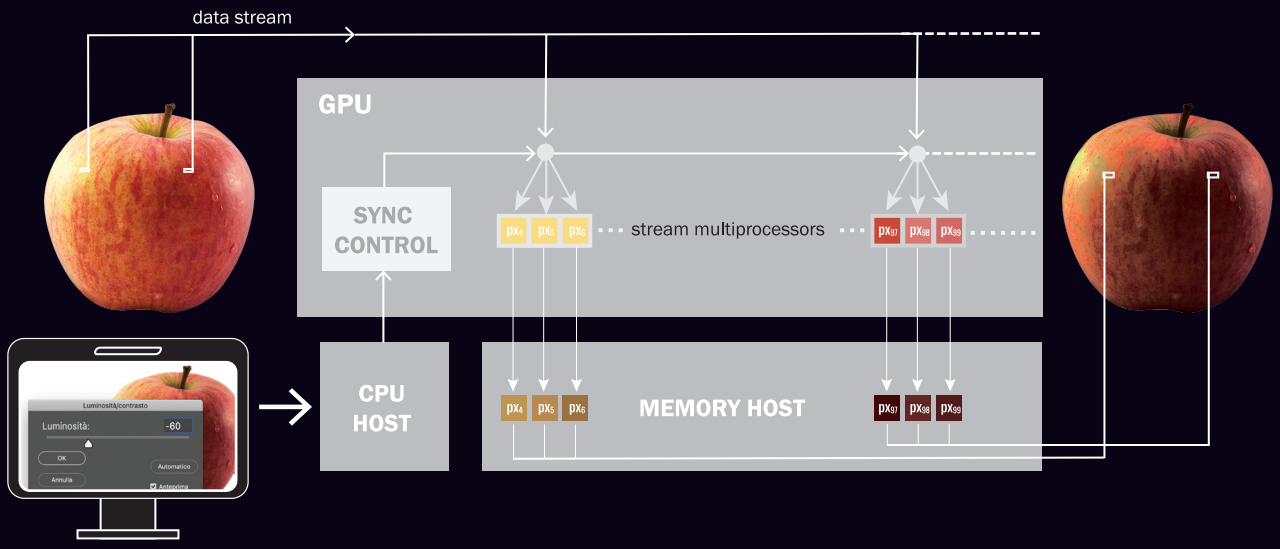
kB chilobyte = 10^3 byte
MB megabyte = 10^6 byte
GB gigabyte = 10^9 byte
TB terabyte = 10^{12} byte
PB petabyte = 10^{15} byte
EB exabyte = 10^{18} byte
ZB zettabyte = 10^{21} byte
YB yottabyte = 10^{24} byte

di calcolo dei grandi esperimenti.
L'Infn ha già fatto passi significativi in questa direzione, stringendo un accordo strategico con il consorzio interuniversitario per il supercalcolo Cineca grazie a cui verrà finanziato "Leonardo", una macchina di calcolo con una capacità grosso modo pari a un quarto di un exascale. Leonardo verrà installato presso il Tecnopolo di Bologna, dove sarà trasferito anche il principale centro di calcolo dell'Infn, il Cnaf (vd. fig. c). In questo modo si realizzerà un grande centro di calcolo unico a livello europeo per potenza complessiva di calcolo e per capacità di memorizzazione e di analisi di dati sperimentali: una struttura fondamentale per il progetto Hilumi-Lhc.
Potenze di calcolo enormi sono e saranno sempre più disponibili anche dai *cloud provider* commerciali con una linea di tendenza dei costi di utilizzo marcatamente al ribasso.
Questo scenario ibrido, dai sistemi privati ai supercomputer e ai sistemi *cloud* commerciali, che disegna la possibilità di un sistema di calcolo a "*plug in*", dove, a seconda del bisogno e della convenienza economica, potrà essere utilizzato uno o più di questi sistemi, sottolinea la necessità di una radicale evoluzione dei modelli di calcolo che includa anche e soprattutto l'evoluzione del software di analisi tracciato qui sopra. Al centro di questo disegno c'è un sistema robusto e veloce di gestione unica dei dati, basato su una decina di grandi "*storage center*" distribuiti a livello planetario, interconnessi tra loro da reti ad altissima velocità e progettati in modo da essere visti come un'unica entità.
Non sono prevedibili, ad oggi, svolte tecnologiche nei prossimi dieci anni tali da sconvolgere le previsioni qui mostrate, anzi si nota un certo rallentamento degli sviluppi industriali dovuti a problemi di saturazione del mercato. La tecnologia del "computer quantistico", che per la sua natura intrinsecamente parallela potrebbe invece avere un effetto dirompente per le nostre applicazioni, è ancora a livello sperimentale. L'interesse però è fortissimo e numerosi progetti di ricerca e sviluppo sono stati proposti dalle comunità, spesso in collaborazione con le industrie del settore.
È comunque interessante notare come la rivoluzione digitale che ci sta portando a una economia e quindi a una società *data driven*, sia progredita e si stia sviluppando quasi contemporaneamente ai grandi esperimenti scientifici di ultima generazione, in particolare di fisica delle alte energie, di astrofisica, ecc., per definizione anch'essi *data driven* e ora capaci di produrre quantità enormi di dati. L'industria informatica sta investendo ingenti capitali in ricerca e sviluppo in questo settore per facilitare decisioni aziendali strategiche e quindi accrescere la competitività delle aziende che utilizzano queste nuove tecnologie. Se sostituiamo la parola "aziende" con "esperimenti", diventa subito evidente come questi sviluppi possano essere di fondamentale supporto alla ricerca scientifica laddove c'è una massiccia produzione di dati. A riprova di ciò va citato il recente accordo tra il Cern e Google, che avrà "l'obiettivo di realizzare progetti congiunti in ambito del *cloud computing*, *machine learning* e *quantum computing*".
Com'è successo per lo sviluppo del web e in innumerevoli altre situazioni, anche in questo caso la scienza ha l'occasione di "adottare" tecnologie già esistenti o in corso di sviluppo per poi rielaborarle inventando per i propri fini una nuova tecnologia che spesso sarà oggetto a sua volta di trasferimento di conoscenza e ricaduta verso il mondo dell'industria e della società.

c.
Il robot del Cnaf (una specie di silos orizzontale lungo 10 m e largo 2,5 m) che gestisce un archivio di 10.000 "*tape*", dove sono memorizzati i dati di Lhc e dei principali esperimenti dell'Infn. Ogni *tape* (in italiano "nastro") è una cassetta magnetica con dimensioni circa il doppio di quelle delle cassette che si utilizzavano molti anni fa per ascoltare musica.

Processori grafici

1.
Rappresentazione schematica del flusso dei dati in una Gpu, in risposta al comando di ridurre la luminosità dell'immagine di una mela.



Le "*graphics processing units*", in breve Gpu, furono progettate principalmente per velocizzare gli effetti 3D da fornire a oggetti che dovevano essere rappresentati sulla superficie 2D dello schermo di un computer, aggiungendo anche effetti visivi (come gli effetti di sorgenti luminose che illuminano l'oggetto, le ombre, la consistenza delle superfici, ecc.) per rendere l'oggetto sempre più simile a quello reale. Questi effetti richiedono numerosi calcoli matematici eseguiti con un alto grado di parallelismo in modo da poter indirizzare velocemente e simultaneamente i milioni di pixel che compongono lo schermo di un computer. La matematica necessaria si limita a trasformazioni geometriche come rotazioni, traslazioni, cambio di sistemi di coordinate, quindi tipicamente delle operazioni tra matrici e tra vettori. Questo limitato set di operazioni da eseguire simultaneamente su più dati permette di architettare un processore specializzato in questo compito e in grado di eseguire la stessa istruzione su più dati contemporaneamente. La figura mostra uno schema semplificato di una Gpu, composto da più "*streaming multiprocessor*", che possono lavorare in parallelo su *stream* di dati diversi. All'interno dello stesso *streaming multiprocessor* agiscono più unità aritmetiche di processamento che eseguono in parallelo lo stesso set di istruzioni sullo *stream* di dati in ingresso. Le moderne Gpu possono avere fino a un centinaio di *stream multiprocessor*, ognuno contenente decine di processori matematici, e quindi complessivamente possono avere migliaia di *core* e possono raggiungere le migliaia di miliardi di operazioni al secondo.

Ogni singolo *stream multiprocessor* ha quindi un'architettura Simd (*Single Instruction Multiple Data*) specializzata in calcolo matriciale e questo estende l'utilizzo delle Gpu al di là della *computer graphics* e dei videogiochi. I due supercalcolatori più potenti al mondo (e altri tre tra i primi dieci) utilizzano Gpu come acceleratori matematici. Inoltre, la capacità di manipolare matrici e, in alcuni esemplari, anche tensori, consente a questi processori di essere usati in molteplici applicazioni che vanno dall'algebra lineare alla statistica, ma soprattutto nei sistemi di intelligenza artificiale come *machine learning* e *deep learning*.

Biografia
Gaetano Maron attualmente è direttore del Cnaf, il Centro Nazionale dell'Infn per la ricerca e lo sviluppo nel campo delle tecnologie informatiche. Collabora all'esperimento di fisica delle alte energie Cms dell'acceleratore Lhc al Cern di Ginevra.

Link sul web
https://www.scientificamerican.com/article/how-close-are-we-really-to-building-a-quantum-computer/

DOI: 10.23801/asimmetrie.2019.27.4